An evaluation of diagnosis-related group severity and complexity refinement

by Thomas E. McGuire

In 1988, an ambitious and extensive project was undertaken in New Jersey to evaluate severity class adjustment of the all-payer prospective payment system. Another project objective was to evaluate alternative strategies for refining diagnosis-related groups (DRGs). The evaluation presented here includes a comparison of DRG refinement using Computerized Severity Index classes and Yale University complexity classes. Statistical methods and payment simulations are used to assess the impact of DRG refinement and consequent revenue changes. When a high volume subset of DRGs is refined, simulated payment shifts between hospitals on the order of 5 percent of total hospital costs are indicated by this analysis.

Introduction

In early 1988, a consortium of large New Jersey hospitals, each with teaching programs, asserted that its case mix was composed of more severely ill and consequently more costly patients for whom the payment system did not provide equitable compensation. This claim implied that the diagnosis-related group (DRG) classification system needed refinement to accurately discern a more costly class of patients within a DRG. The consortium contended that payments should reflect patient severity levels within DRGs.

In response to this and other claims of payment inequities, the New Jersey Department of Health (NJDOH) and 25 New Jersey hospitals initiated an evaluation of DRG refinement based on the Computerized Severity Index (CSI) system. The strategy was to use CSI to assign severity levels to patient records from 1986 and then to form patient classes within DRGs based on the severity levels. Severity class cost weights within DRGs would then be determined and, if the evidence supported refinement, the payment system would be modified to reflect severity cost differences. The project began in earnest in May 1988; data collection was completed by January 1989. Analysis and payment simulations continued throughout the summer of 1989, and the project final report was delivered in November 1989.

The State of New Jersey was required to modify its all-payer payment system when it was no longer permitted to include Medicare patients in its all-payer policy. In addition to the considered modification of DRGs by severity classes, policy revision included implementation of the New York (NY) State modified DRG classification (NY 6.0 DRGs) and revised DRG base payments.[1] A new strategy to account for differing costs among hospitals as a function of teaching intensity was also included in the payment policy revision.

The principal objective of the evaluation project was to determine if CSI patient severity of illness could identify case-mix differences between New Jersey hospitals that correlated to higher patient care costs. An important expansion of the study undertaken during the analysis phase was the evaluation of the Yale University DRG refinement. Coordination and operational support of the study were performed by Health Systems International, the developer and vendor of the CSI software system, and Quantitative Consultants, Inc., the developers of the CSI severity methodology (Horn, 1986).

The analysis phase of the project focused on identifying DRGs showing the greatest improvement in homogeneity after refinement and consequent payment simulations to determine financial impact. Both severity and complexity refinements are based on a more specific use of secondary diagnoses than are currently used in Medicare DRGs. Complexity refinement depends on the existence of specific secondary diagnoses, whereas the severity refinement uses secondary diagnoses to motivate evaluation of more detailed clinical information. Each refinement uses a class structure that reflects different levels of treatment intensity within each DRG. Multiple classes are formed for each DRG, thereby forming a class structure overlay on the DRGs. Although the principal DRG refinement models compared and contrasted in the evaluation were the CSI severity refinement and Yale University complexity refinement, a number of other DRG class models were investigated. Revelations from the model analysis led to a significant new refinement of the DRGs used in New York and New Jersey.

Severity and complexity methods

CSI is a disease-based system that uses coded diagnosis information from a patient record to motivate queries for clinical data that are then used to evaluate patient severity of illness. CSI refinement of DRGs consists of using the four-level severity score to create four severity classes within each DRG.

Clinical data required to implement CSI consist of laboratory test results, vital signs, medical history, and physical condition factors. Therefore, assigning a severity class requires information not normally present on the discharge abstract but found only within patient clinical records. The lowest level of the four-level patient severity score corresponds to clinical findings in the normal range

---

[1] The New York State Department of Health adopted a modification of Medicare DRG definitions to improve classification of the general patient population. New York DRGs contain a new major diagnosis category (MDC) for patients with the human immunodeficiency virus, a new MDC for multiple trauma patients, a restructured newborn MDC based on birth weight, a restructured drug and alcohol abuse MDC, modified DRGs for cystic fibrosis, and several new DRGs for treatments such as liver and bone marrow transplants. The New York State payment system also requires that DRG inlier patients meet both low and high DRG length-of-stay thresholds.

and is indicative of low-risk patients. The highest level corresponds to patients who have serious clinical findings that indicate life threatening diseases.

Evaluating the severity of each disease present for a patient is the first step in determining the patient's overall severity level. The severity of each disease is also measured on a four-level scale (mild, moderate, severe, and life threatening). Patient diseases are indicated by the diagnosis information attested to by the attending physicians. Each patient-specific diagnosis is assigned to a disease group, and the severity of the disease group is evaluated. Within CSI, each disease group is indicated by a set of diagnosis codes from the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM). A disease group may include a patient's principal diagnosis and/or secondary diagnoses. The patient's overall severity level is determined from the levels and interactions of all disease groups present. The principal disease group—not necessarily indicated by the principal diagnosis—yields the greatest contribution to the overall severity score, but serious secondary diseases or secondary diseases that strongly interact with the principal disease may also have considerable impact on the overall score.[2]

Severity of illness can be measured over different time periods during a patient's hospital stay. Two valuable periods of time for which severity is measured are the admission period and the entire period of hospitalization. The admission period is typically considered to be the first 48 hours of the patient's stay, and the associated severity level is called the "patient's admission severity level." The maximum severity level represents the highest severity level that could have been attained during the patient's stay. The maximum severity level uses patient data acquired during the entire hospital stay; it was chosen for the retrospective evaluation project.

The Yale University complexity refinement was developed during 1986-88 by the Health Systems Management Group, supported by a Health Care Financing Administration (HCFA) grant (Fetter et al., 1989). A class structure overlay on Medicare DRGs was created by classifying diagnoses coded as comorbidities or complications (CCs). Information required for patient class assignment using this class structure overlay is the same as that required for DRG assignment and is available in standard patient discharge abstracts. Consequently, patient clinical records are not needed.

Structural modifications to the Medicare DRGs began with a recombination of adjacent diagnosis-related groups (ADRGs) that were previously distinguished by the existence of specific CCs.[3] Recombination reduced the number of DRGs from 475 to 318. The next step was the creation of a class overlay that partitioned each ADRG into multiple subgroups. The class structure overlay is based on ADRG-specific CC class assignments that were determined from extensive statistical and clinical evaluation. Class levels are sequentially ordered with

respect to increasing expected resource (cost and length of stay (LOS)) requirements. Three classes are defined for patients in medical ADRGs and four classes for patients in surgical ADRGs. The overall ADRG class assignment for a patient is the maximum class level assigned among the patient's CCs. The class of an ADRG is therefore derived from the class assignment of one of the CCs that is coded on a patient abstract.

In summary, the principal differences in class assignment methodology between CSI and Yale are CSI's use of detailed patient clinical information and its processing of interactions between the principal and secondary diseases. Severity or complexity classes can be applied to either DRGs or ADRGs that contain CC information, enabling direct comparison of Yale and CSI class structures. Two other significant changes recommended by the Yale study were the identification of extreme low-cost medical patients (class L) and extreme high-cost tracheostomy patients (class H). These classes are not included in the analysis because they are specifically accounted for in the New Jersey DRG system. The low-cost medical patients are those that expire in the first 48 hours of hospitalization; they are almost always excluded by the low length-of-stay trim points used in New Jersey. All patients having tracheostomy procedures are grouped into two DRGs, 735 and 736, in the New York State modified DRGs used in New Jersey.

Study hospital characteristics

How may the results of this study be extended to areas outside New Jersey? The study hospitals appear to be a representative metropolitan sample and the severity impact results are likely to be indicative of the impact other metropolitan area hospitals would experience.

There were 94 nongovernment, not-for-profit community hospitals operating in New Jersey in 1986, the year from which the study data were selected. Those hospitals had 1.05 million admissions during 1986 and employed over 60,000 full-time personnel. The operating expenses for the hospitals in 1986 were nearly $4 billion, or approximately $3,000 per admission or $430 per inpatient day (American Hospital Association, 1987).

New Jersey does not have any hospitals classified as nonmetropolitan. In the entire United States, 46 percent of hospitals are considered nonmetropolitan. However, nonmetropolitan hospitals account only for 23 percent of beds, 20 percent of patient days, and 2 percent of full-time equivalent (FTE) residents. New Jersey and U.S. hospital size distributions are also noticeably different. The percent of small hospitals (less than 100 beds) in New Jersey is about 5 percent, whereas it is about 50 percent for the Nation overall. Small rural hospitals account for one-half of all hospitals in our country and are not well represented in the hospital distribution of New Jersey. Yet, over 20 percent of admissions, expenses, surgeries, emergency room visits, and births occur in large hospitals (more than 500 beds) for both New Jersey and the United States. Thus, New Jersey hospitals reflect the country's demographics with respect to the concentration of population and services in metropolitan areas.

Table 1

Percent distribution of study hospitals, New Jersey hospitals, and national hospitals, by selected characteristics

Characteristic	Study	New Jersey	United States
Hospital size		Percent distribution	
Less than 200 beds	8	23	70
200-500 beds	48	62	25
500 beds or more	44	15	5
Full-time equivalent residents			
Less than 200 beds	—	4	4
200-500 beds	—	44	32
500 beds or more	—	52	64
Teaching status			
Nonteaching	28	64	—
Other teaching	16	14	—
Teaching	56	21	—
Hospital location			
Inner city	36	23	—
Urban	24	22	—
Suburban	28	39	—
Rural	12	16	—

SOURCE: American Hospital Association (1987); and 1986 New Jersey Department of Health hospital files.

How representative were the study hospitals of all New Jersey hospitals? The study hospitals were more representative of location than of size with respect to all State hospitals and contained a higher percent of hospitals with physician training programs than the State population, according to the peer group classification used in New Jersey.

All hospitals in New Jersey were invited to participate in the Severity of Illness Evaluation Project and, as a result of a statewide meeting between the NJDOH and the hospital managers, 25 hospitals volunteered to participate. The participants represented 28 percent of New Jersey hospitals and serve 34 percent of New Jersey patients. A comparative summary of all national hospitals, all New Jersey hospitals, and the study hospitals is given in Table 1.

The peer group classification of hospitals for degree of teaching programs may not be a good indicator of teaching program impact, and it should be noted that during 1988 an investigation characterized the impact of physician training programs on New Jersey hospital expenses (Network Inc., 1989). One consequence of that investigation was the application of a set of factors to adjust hospital DRG payments for the indirect expense of supporting physician training programs. Adjustments for the study hospitals appear to be a uniform random sample of the adjustments for all State hospitals.

Outline of the evaluation project

A brief overview of the project organization and process is presented in this section; further detail is available in the project final report (Health Systems International, 1989). Information flow and requisite tasks are illustrated in Figure 1. The two bold-faced boxes contain the obligations of the NJDOH and the New Jersey hospitals. The other boxes represent the processing undertaken by staff at Health Systems International (HSI). Details of the analysis data base construction are presented in the following section.

Based on a clinical review of each DRG, a preliminary list of DRGs to be included in the study was developed. The focus of the study was on medical and surgical patients—psychiatric, substance abuse, and elective surgery patients were eliminated. In general, elective surgery is only performed on patients at low severity levels and, therefore, there are few high-severity patients in elective surgery DRGs. Pediatric patients were excluded because CSI severity criteria for pediatric patients were not available. In addition, the DRG frequency distribution of patients was used to exclude DRGs with insufficient patient volume in the study hospitals. Available data on the relationship between severity and cost by DRG were also reviewed before selecting the final DRGs for the study. A total of 131 DRGs, representing approximately 40 percent of New Jersey hospital patients, were excluded.

Following removal of patient records within excluded DRGs, patient records were randomly selected from the 1986 NJDOH hospital files.[4] The random sample was formed at the hospital level so that each hospital would have a proportionate share of records to rate; most hospitals had approximately 4,000 patient records to rate.

To ensure that each hospital used accurate information and focused on severity issues during the study, CSI severity worksheets containing patient-specific clinical questions were generated for each hospital's selected records. A clinical analysis report was also produced for each patient abstract record to point out clinical inconsistencies in the records so that abstract and clinical chart information could be checked prior to completing the worksheet.[5] Delivery of worksheets to each hospital was synchronized with hospital staff training. Within 3 weeks of training, a visit was made to each hospital by a

---

[4]The 1986 hospital files represented the most current files containing audited cost data.
[5]Inconsistent diagnoses and procedure data were detected using the HSI Clinical Data Editor. Error checks included age, sex, and discharge status problems as well.

Figure 1
New Jersey Severity of Illness Evaluation Project

```
                          ┌─────────────────┐
                          │    NJ DOH:      │
                          │     1986        │
                          │  audited cost   │
                          │      and        │
                          │  abstract data  │
                          └─────────────────┘
          ┌────────────────────┐      ┌────────────────────────┐
          │ DRG selection and  │      │ Assign New York 6.0 DRGs│
          │random sampling of  │      │ and determine LOS trim  │
          │     abstracts      │      │        points           │
          └────────────────────┘      └────────────────────────┘
                   │                             │
                   ▼                             ▼
          ┌────────────────────┐      ┌────────────────────────┐
          │Produce worksheets  │      │                        │
          │       and          │      │  Create cost data bases│
          │abstract quality    │      │                        │
          │     reports        │      │                        │
          └────────────────────┘      └────────────────────────┘
                   │                             │
                   ▼                             ▼
          ┌────────────────────┐      ┌────────────────────────┐
          │New Jersey hospitals:│     │Population statistics and│
          │Fill in worksheets, │      │   sample fractions      │
          │data entry,         │      │                        │
          │and archive data    │      │                        │
          └────────────────────┘      └────────────────────────┘
                   │                             │
                   └──────┐           ┌──────────┘
                          ▼           ▼
                     ┌──────────────────────────┐
                     │ Create analysis data base,│
                     │ perform model analysis, and│
                     │   payment simulations      │
                     └──────────────────────────┘
```

NOTES: NJ DOH is New Jersey Department of Health. DRG is diagnosis-related group. LOS is length of stay.

SOURCE: (McGuire, 1989.)

nurse or clinician to evaluate the reliability of severity raters and offer suggestions on efficient practices for the rating process.

Financial reports submitted by the hospitals after completion of the study indicated that the cost borne by the hospitals was approximately $20 per record.

Analysis data base

An analysis data base containing patient demographics, clinical severity, and audited hospital costs was constructed for modeling alternative patient class assignment methods and payment scenarios. The clinically edited patient abstract data, audited direct hospital costs, and severity scores made this data base unique for payment analysis.

Data base fields were defined for inlier or outlier status, hypothetical payment, complication and comorbidity indicators, diagnosis-specific Yale complexity refinement, and CSI classes. The analysis data base included patient information required for DRG assignment: age, sex, discharge status, diagnoses, and procedures. Other data base fields identified the payer, reported charges, LOS, and attending physician. Audited information acquired from the NJDOH identified direct hospital costs and acute LOS for each record and

hospital-specific payment rates used in the New Jersey all-payer PPS. The cost data used were direct patient-care cost and excluded the cost of capital, certain administrative costs, and staff physician costs. The direct patient-care costs were computed using NJDOH cost-to-charge ratios to convert charges into cost. The direct patient-care costs for each hospital were equalized for wage-rate variations and for differences in the indirect cost of teaching programs. Acute LOS does not include the number of days a patient may have been awaiting transfer while no longer requiring hospital care. This NJDOH information was merged with the severity information collected by the hospitals and was used to determine sample-to-population ratios for each hospital and DRG combination.

During construction of the data base, each patient record was classified using three DRG classifications: Health Care Financing Administration (HCFA) version 6.0 DRGs, NY version 6.0 DRGs, and Yale refinement of the HCFA version 6.0 DRGs. Data base fields were included to indicate errors or conflicts for each diagnosis and for other problems in the record, such as problems related to clinically inconsistent interactions between diagnoses, procedures, age, sex, and LOS.

Outlier status fields were used to indicate low and high LOS or cost outliers. Different policies for defining

Table 2
Distribution of severity levels for study diagnosis-related groups

Refinement model	Severity class 1	2	3	4
Computerized Severity Index				
Percent of count	69	17	6	8
Count	52,656	12,603	4,340	6,306
Percent of cost	45	22	11	23
Average cost	$ 1,462	$2,960	$4,327	$6,266
Yale refined class				
Percent of count	43	37	16	4
Count	32,606	28,125	12,236	2,938
Percent of cost	24	32	29	15
Average cost	$ 1,287	$1,958	$4,083	$8,709

SOURCE: Calculated from 1988 New Jersey Severity of Illness Evaluation Project Data.

outlier status were used in the analysis. One policy used the low- and high-trim points defined for payment purposes by the NJDOH. Currently, New York and New Jersey use low LOS trim points and per diem payments for short-stay patients. The trim points are influenced by policy considerations and thereby can confound statistical evaluation. The statistical outlier policy used in this analysis was based on an algorithm for robust estimation of central tendency.[6] A nonparametric trimming algorithm is important when evaluating homogeneity properties of different classification systems, because parametric methods require knowledge of the underlying data distribution. DRG cost and charge distributions have a lognormal shape, and that tendency holds for Yale within-DRG complexity classes as well. The within-DRG CSI severity distributions tend to have normal distributions. The nonparametric trimming algorithm defines both low- and high-trim values based on the first and third quartiles (Q1, Q3) of the DRG data distribution. The formulas defined for reasonably symmetric distributions are:

$$\text{low} = [Q1 - 1.5 * (Q3 - Q1)] \text{ and high}$$
$$= [Q3 + 1.5 * (Q3 - Q1)].$$

Given that DRG data are lognormally distributed, one should logarithmically transform such data before calculating trim points. An alternative to transforming the data is to adjust the above formulas to the following:

$$\text{low} = [Q1/(Q3/Q1)^{1.5}] \text{ and high} = [Q3 * (Q3/Q1)^{1.5}].[7]$$

The low trim point was forced to be nonnegative if calculation yielded a negative value. Data within a DRG were trimmed before classes were formed and models compared.

A summary of 75,905 untrimmed records, each of which is rated in the range 1-4 by both systems, is presented to show crude similarities and differences.[8] The comparison is crude because class relations are not invariant across DRGs. For example, the relation between

Class 1 and Class 2 patients is not uniform within all DRGs for either system. Also, for medical records there are at most three Yale classes. Given these caveats, the distribution of severity scores and average cost for DRGs in the study is shown in Table 2. Several properties of the data in Table 2 are representative of equivalent reports on individual ADRGs. CSI tends to rate a larger number of patients in the level 1 class than does Yale. The size of the fourth CSI class tends to be larger than the Yale class because of the CSI level 4 death rating. The inclusion of deaths in CSI level 4 lowers its average cost. DRGs currently distinguish one group of patients (DRG 123—Circulatory disorders with acute myocardial infarction, expired) based on whether the patient died during the hospitalization. The use of an early medical death class, or rating deaths in the most severe class, extends the use of this controversial variable within the payment system. The incentives associated with such payment decisions should be considered carefully.

The model report shown in Table 3 is a summary of Yale and CSI class refinements for ADRG 148—major small and large bowel procedures. Model reports summarize the effects of class formation across three dimensions—LOS, cost, and frequency. The cost differences between severity classes are indicated by average cost and the ratio of class average cost and ADRG average cost—the class cost weight. Variance changes resulting from class formation are summarized at the bottom of each model report.

The model report displays similarities and differences between the CSI and Yale class structures that occur for many other ADRGs as well. For example, one similarity is the relationship between different class averages; higher class levels typically have higher average cost and LOS. There are ADRGs with exceptions to that observation; and within this data set, they occur more frequently for Yale classes than for CSI classes. A typical difference in the class structures is the range of class average costs and resulting cost variance reduction; CSI cost weights have a larger range and yield considerably better cost variance reduction, 31.31 percent versus 16.46 percent for Yale.

Class model analysis

Various class model designs were created to explore questions about the impact on payment policy resulting

---

[6]"Multiply-Skipped Mean, Max(3k, 2) Deleted" (Andrews, Bickel, Hampel, et al., 1972)

[7]These formulas follow from the observation that the log transform is a monotonic transform that will not change the quartile locations.

[8]For the purposes of this untrimmed data report, Yale class L (early medical death) patients were rated as level 1. CSI rates all deaths as level 4.

Table 3
Computerized Severity Index and Yale model reports

ADRG# 148 Major small and large bowel procedures
DRG 148—New Jersey State Inlier Count: 7,048   Low LOS Trim: 4   High LOS Trim: 52
DRG 149—New Jersey State Inlier Count: 1,705   Low LOS Trim: 3   High LOS Trim: 23

Model: CSI

Class	#	PCT	ALOS	SDLOS	CVLOS	ACOST	COSTW	SDCOST	CVCOST
1	251	32.14	12.06	5.26	0.44	3,518.66	0.62	1,647.48	0.47
2	285	36.49	16.51	8.11	0.49	4,958.82	0.88	2,338.51	0.47
3	114	14.60	19.30	8.98	0.47	6,288.68	1.12	4,074.71	0.65
4	131	16.77	23.66	12.81	0.54	10,601.96	1.88	6,703.02	0.63
ADRG	781	100.00	16.68	9.38	0.56	5,636.64		4,309.18	0.76

	%RED VAR	%RED MAD	%RED CV	WCV
COST	31.31	22.25	31.17	0.52
LOS	18.27	12.47	14.39	0.48

Model: Yale

Class	#	PCT	ALOS	SDLOS	CVLOS	ACOST	COSTW	SDCOST	CVCOST
1	84	10.76	10.96	3.98	0.36	3,149.81	0.56	1,132.96	0.36
2	92	11.78	13.09	6.17	0.47	3,825.87	0.68	1,827.75	0.48
3	344	44.05	16.10	8.20	0.51	4,965.56	0.88	2,813.62	0.57
4	261	33.42	20.56	11.22	0.55	7,959.75	1.41	5,874.76	0.74
ADRG	781	100.00	16.68	9.38	0.56	5,636.64		4,309.18	0.76

	%RED VAR	% RED MAD	%RED CV	WCV
COST	16.46	12.71	21.89	0.59
LOS	11.64	8.43	10.26	0.50

NOTES: DRG is diagnosis-related group. # and PCT are the number and relative percent of records.
ALOS, SDLOS, and CVLOS are average, standard deviation, and coefficient of variation of length of stay, respectively.
ACOST, SDCOST, and CVCOST are similarly defined for cost.
COSTW is the class average cost divided by the adjacent diagnosis-related group average cost.
ADRG are aggregate statistics for the adjacent diagnosis-related group.
WCV is the weighted coefficient of variation—class CV weighted by fraction of cases (PCT).
N_s is the number of data and CV_s is the coefficient of variation for cost in class s.

$$WCV = \sum_s \frac{N_s}{N} CV_s \qquad \% \ REDCV = 100 \cdot [1 - WCV \div CV]$$

% RED VAR, % RED MAD, and %RED CV are the reduction in variance, mean absolute deviation, and coefficient of variation, respectively $C_{i,s}$ is the datum cost; C_s is class s average cost; and C is DRG average cost.

$$\% \ REDVAR = 100 \cdot \left[1 - \sum_{i,s} (C_{i,s} - C_s)^2 \div \sum_{i,s} (C_{i,s} - C)^2 \right]$$

$$\% \ REDMAD = 100 \cdot \left[1 - \sum_{i,s} |C_{i,s} - C_s| \div \sum_{i,s} |C_{i,s} - C| \right]$$
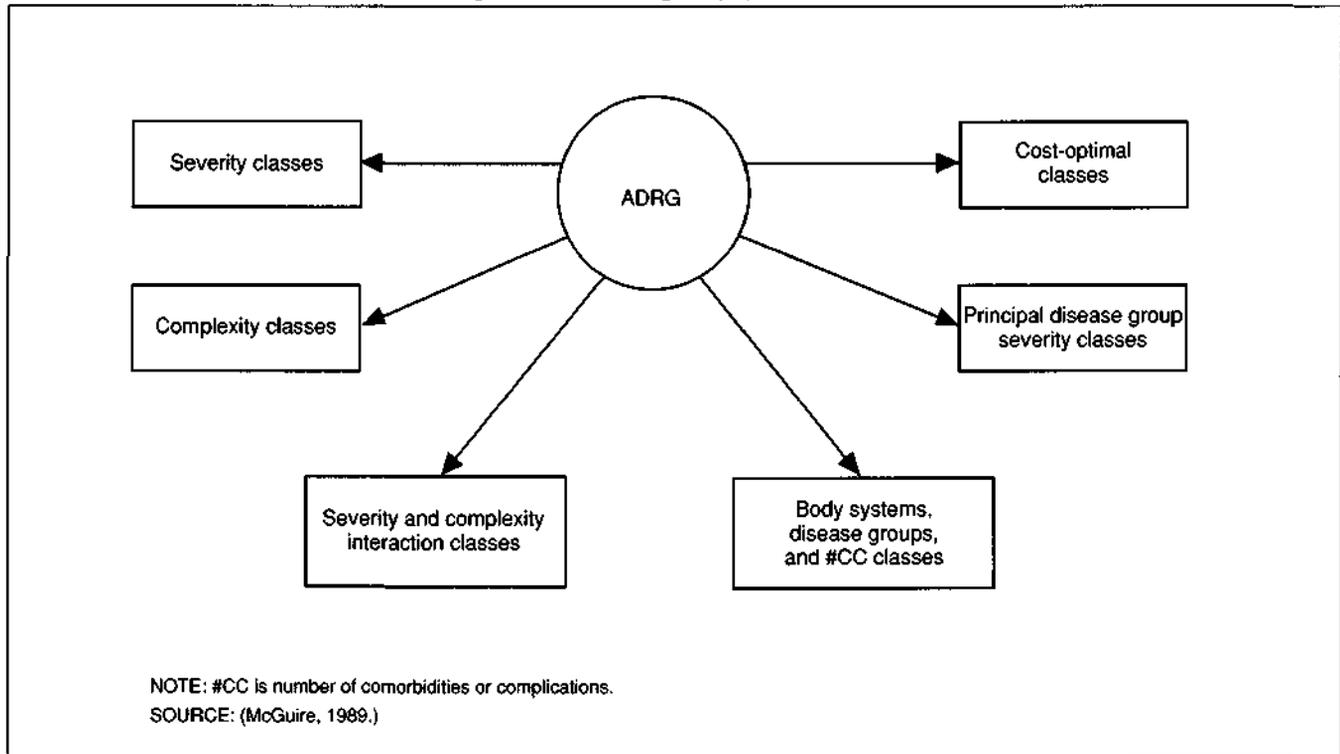
SOURCE: Calculated from 1986 New Jersey Severity of Illness Evaluation Project data.

from the formation of patient classes within DRGs. As previously described, severity and complexity models were formed using the CSI and Yale class definitions. Other comparative models were constructed to evaluate interactions between complexity and severity and to test alternative CC-derived complexity measures.

The models used can be functionally grouped into the categories displayed in Figure 2. The model groups that will be described in this section are severity classes, complexity classes, body system and disease groups, and cost optimal classes. A fundamental issue regarding any class overlay on DRGs is the number of classes into which each DRG should be subdivided. Creating too many classes leads to difficulties with predictive analysis and policy formation as the number of patients in a class is likely to be too small for reliable estimates of resource requirements. Also, it is unlikely that the same number of

classes should be created for every DRG because the number of patients in different DRGs varies widely. Adjacent class combinations were modeled for both Yale and CSI classes to evaluate the significance and appropriateness of the number of defined DRG classes. The analysis data base contains six models for all possible CSI adjacent class recombinations within a four-class structure. There are similarly six such models for Yale adjacent class recombinations. The class recombination models are summarized in Table 4, where the letters "C" and "Y" represent CSI and Yale class recombinations, respectively. The number of classes used in a model is indicated by the highest number in the mapping. For example, "Y(1,2,2,3)" indicates that a four-class Yale DRG refinement was re-formed into three classes in which the second- and third-class patient records were merged. This results in three patient classes

Figure 2

Adjacent diagnosis-related group (ADRG) class models

NOTE: #CC is number of comorbidities or complications.

SOURCE: (McGuire, 1989.)

with no change in the first or fourth Yale classes but with the second and third Yale classes merged into a single class labeled "Class 2;" the fourth class was subsequently relabeled.

The relative model results, given in Table 4, are similar for most ADRGs. For all merit measures, CSI class refinement (Models 1-7) dominates the Yale class refinement (Models 8-14). Model 15 is the status quo, whereby the existence of CCs determines DRG class (e.g., DRGs 148 and 149). Both Yale and CSI classes are a considerable improvement over Model 15. The three-class models, based on specific adjacent class merging, CSI (Models 2,6,7) and Yale (Models 9,13,14), have comparable merit measures to the four-class models (Models 1,8).

Other class models were created to explicitly test the interaction of inhospital death and severity class with payment policy. In the four-class CSI system, all patients who die in the hospital are classified as overall severity level 4. However, Yale refinement research indicated that early death in medical DRGs has a high association with low cost. A five-class model that distinguished all deaths as a separate class was examined.

The existence of complications and/or comorbidities on the abstract record does not necessarily indicate that more costly care was provided. Another set of models enabled a test of the hypothesis that there were CCs on the abstract record that had no supporting clinical evidence, as evidenced by CSI severity evaluation, and therefore should not be used for class assignment.

Models were formed to test the interaction of cost with the existence and number of CCs, the number of CCs in distinct body systems—such as major diagnosis categories (MDCs)—and the number of CCs in distinct disease groups. These models were used to test the hypothesis that the number of CCs is a more powerful indicator of high cost than just the existence of CCs.

Optimal class assignment models were used to form a basis for overall comparison. Optimal class assignments were used to determine the maximum possible explanatory power of two, three, or four classes within a DRG. Rather surprising results were realized from the optimal-cost class models.[9]

Models 16 and 20, with five classes (CSI with death class and CSI with CC split), achieve the highest cost-variance reduction. The fifth class in Model 16 distinguishes patients that expire during their hospital stay. Thus, the improvement in Model 16 over Model 1 resulted solely from moving some of the Class 4 patients into Class 5. The fifth class in Model 20 contains cases that have no CCs (e.g., the original cases in DRG 149). The other four classes in Model 20 correspond to the four CSI severity levels assigned to patients having CCs (e.g., the original cases in DRG 148). Whereas Model 16 distinguishes the most severe patients, Model 20 attempts to distinguish the least severe patients and does not require a severity score for patients without CCs.

---

[9]McGuire and Bender (1990) describe the specific algorithms used to determine optimum DRG variance reduction using class structures. The algorithms are also detailed in McGuire (1989a).

Table 4
Summary of model reports for adjacent diagnosis-related group (ADRG) 148

Model	#	%Red VARCost	% Red CVCost	% Red MADCost	% Red VARLOS	% Red CVLOS	% Red MADLOS
C(1,2,3,4)	1	31.31	31.17	22.25	18.27	14.39	12.47
C(1,2,2,3)	2	30.32	28.58	21.33	17.35	13.87	11.51
C(1,1,2,2)	3	21.57	25.05	15.12	12.73	9.03	8.92
C(1,2,2,2)	4	11.46	15.83	11.17	11.55	10.52	8.52
C(1,1,1,2)	5	26.79	22.58	15.71	11.16	8.03	6.32
C(1,2,3,3)	6	23.48	28.66	18.63	16.59	13.13	12.19
C(1,1,2,3)	7	29.40	27.56	18.74	14.41	10.30	9.20
Y(1,2,3,4)	8	16.46	21.89	12.71	11.64	10.26	8.43
Y(1,2,2,3)	9	15.81	20.50	11.80	10.68	9.42	7.74
Y(1,1,2,2)	10	7.14	11.23	6.45	7.04	6.21	5.49
Y(1,2,2,2)	11	4.02	6.84	3.73	4.49	5.43	3.71
Y(1,1,1,2)	12	14.61	17.52	10.21	8.60	5.35	5.79
Y(1,2,3,3)	13	7.28	12.04	6.73	7.33	7.12	5.76
Y(1,1,2,3)	14	16.33	21.08	12.43	11.35	9.36	8.17
CC Split	15	3.94	6.77	3.66	4.39	5.39	3.63
CSI 5	16	33.91	32.46	22.74	19.72	14.40	13.25
Class CC Count	17	14.54	20.17	11.52	12.48	9.06	8.82
NDXMDC	18	14.38	18.86	11.46	11.10	6.44	7.89
NDXGID	19	14.73	18.58	11.87	12.33	8.04	8.79
CSI on CC	20	31.89	32.78	22.94	19.31	15.95	13.49
Optimal2	21	61.97	43.63	34.25	35.80	19.62	18.78
Optimal3	22	80.13	62.91	55.91	51.61	35.32	32.09
Optimal4	23	87.33	71.46	65.39	57.65	43.59	37.47

NOTES: %RED VARCost, %RED CVCost, and %RED MADCost are the reduction in variance, coefficient of variation, and mean absolute deviation, respectively, calculated using ADRG 148 cost data from all study hospitals. Likewise, the values calculated using length-of-stay data are labeled %RED VARLOS, %RED CVLOS, and %RED MADLOS. The formulas used are those cited in the notes of Table 3. The model labeled C(1,2,3,4) is a four-class CSI severity partition of ADRG 148 patient cases; Y(1,2,3,4) is the Yale refined complexity partition of the same cases. Models labeled with repeated values show the result of collapsing adjacent partitions into a single partition, such as the CSI three-class partition (Classes 2 and 3 collapsed) labeled C(1,2,2,3). The models are further described in the text.

SOURCE: Calculated from 1988 New Jersey Severity Evaluation Project data.

Models 15, 17, 18, and 19 are models based on counting CCs. Model 15 is the current CC split criteria, as described earlier. Model 17 is a straight forward refinement with a view towards considering interactions of CCs; there are four classes formed based on the number of CCs. Model 18 is a refinement of the counting method wherein the count is the number of CCs in different MDCs, a surrogate for different body systems. Model 19 classes are based on the number of CCs in different CSI disease groups.

Further research into simple counting models based on significant CCs—perhaps as given by the Yale refinement CC classes—may show that carefully chosen CC or disease interaction lists can achieve considerable improvement. The motivation behind Models 17-19 was the hypothesis that, as in combinatorial analysis, the transition from two to three options generally is the transition from manageable problems to particularly challenging ones.

The cost-optimal Models, 21-23, use two, three, and four classes to explain the maximum cost variance possible. Surprisingly, research into these models has shown that, with only two cost-optimal classes, a cost-variance reduction of 60 percent is typical (McGuire, 1989a). Three and four optimal classes typically achieve 75 percent and 90 percent cost-variance reduction. For ADRG 148, the five-class model (Model 16) had the highest cost-variance reduction (33.91 percent), whereas the two-class optimal model (Model 21) achieved 61.97 percent cost-variance reduction. Cost-optimal classes also reveal that approximately 10 percent of cost variance within an ADRG cannot be explained by any

four-class overlay. Thus, a measure of the cost-variance noise or inherent variation within a four-class DRG model is on the order of 10 percent.
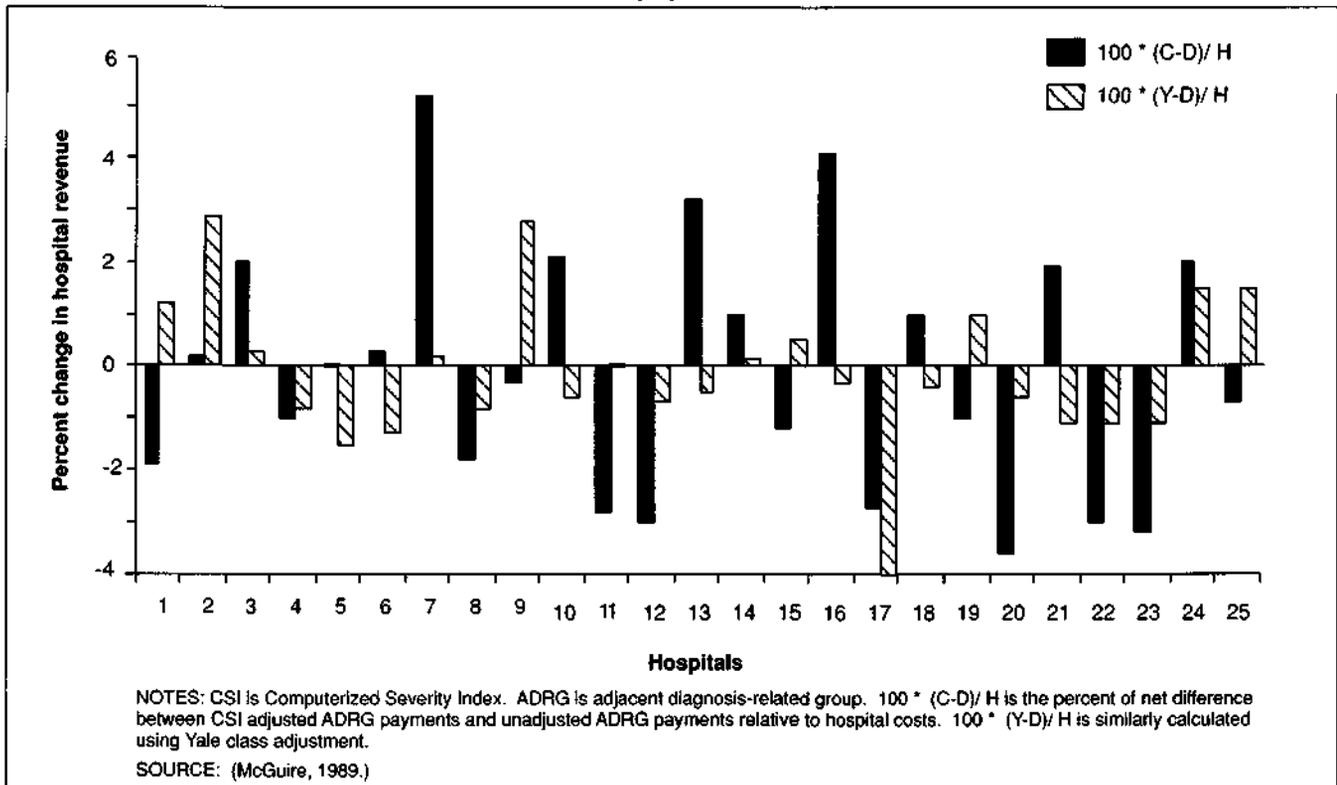
Empirical results using select models within a subset of ADRGs are presented following the next section.

Diagnosis-related group payment evaluation

The previous section described different class overlay models. In this section, consideration is given to payment equity resulting from making the Yale and CSI class-adjusted DRG payments. Payment equity was investigated by simulating hospital payments using two scenarios. The first scenario simulated payment by using the status quo; the second scenario used class cost weights to modify the status quo. The simulations are based on a subset of DRGs. After a short discussion of the subset data used, the cost and payment results are described.

The DRGs selected spanned the full range of MDCs, were manageable in number, represented a significant portion of the hospital costs, and had a measurable effect on variance reduction. The selection process used several criteria. First, selection was made from the New Jersey study data base. Second, the DRGs in each MDC were ranked according to total reported cost within the 25 hospitals. This ranking used population statistics from the NJDOH. The top 10 DRGs in each MDC were picked as candidate DRGs. Third, the Yale and CSI class model reports were examined for each candidate DRG; if a model report showed a cost-variance reduction greater

Simulated class payment revenue shifts

NOTES: CSI is Computerized Severity Index. ADRG is adjacent diagnosis-related group. 100 * (C-D)/ H is the percent of net difference between CSI adjusted ADRG payments and unadjusted ADRG payments relative to hospital costs. 100 * (Y-D)/ H is similarly calculated using Yale class adjustment.
SOURCE: (McGuire, 1989.)

than 5 percent, the candidate was retained. The identified subset of DRGs accounted for 40 percent of the cost in the 25 hospitals and 30 percent of the patients in 1986.[10]

Given the DRG subset, the case data used in the simulation was selected. Because the cases are used to make inferences for overall impact, care was taken to avoid erroneous inferences. In the inference procedure, the sample data from each hospital are used to create fractions of patients within each DRG class at the particular hospital. For population inferences, these fractions are applied to the total DRG population for the hospital to arrive at overall class frequencies. Hospital-DRG combinations were selected if there were at least 10 cases in the hospital-DRG population and, further, if the sample proportion was at least 20 percent of the population. Therefore, each hospital-DRG combination within the data set had to have at least two cases.

After the restriction of hospital-DRG combinations, approximately 30,000 cases were used to used create class fractions. There were an average of 63 DRGs per hospital and an average of 15 hospitals per DRG. There were 103 DRGs selected that were recombined into 68 ADRGs. Of the 25 most frequently occurring DRGs nationwide, 20 are included in the data set (American College of Surgeons, 1988).

To simulate the impact of classes on aggregate payments to the hospitals, the 1986 number of patients for each of the 25 hospitals within the ADRGs was

obtained from the NJDOH payment report. A payment amount for each ADRG was determined based on the average ADRG cost for all hospitals by using wage and teaching equalized costs as determined by the NJDOH. For each hospital and ADRG, the actual number of patients in 1986 was split proportionally into a class-level number of patients based on the class percent determined from the study data base. For each ADRG, the standard payment was scaled into class payments by multiplying the standard payment by the class cost weight determined from the study data base.

The impact on aggregate hospital payments was simulated by comparing the payment a hospital would have received if it were paid the ADRG payment with the ADRG payment scaled by the appropriate class cost weight. The ADRG payment was the average cost for all patient records in the ADRG. The class cost weight was the average cost for all patient records in the particular class divided by the ADRG average cost. Because the ADRG payments were based on the average cost for all patients, the simulation is zero-sum; the excess that one hospital receives must be made up through the shortfalls of other hospitals. The impact of differences between class-refined payments and ADRG payments is evident in Figure 3, where the net differences are expressed relative to the aggregate hospital costs. The differences in payment between refined payments and ADRG payments is on the order of −4 percent to +6 percent. New Jersey

---

[10]During 1986, there were 329,695 inlier patients in the 25 study hospitals. The total reported cost for these patients was $678,340,999 (as calculated from the 1986 New Jersey data set).

Table 5
Cumulative cost and count for the top 25 of the selected diagnosis-related groups

DRG	Description	Percent of cumulative cost	Percent of cumulative count
127	Heart failure & shock	2.3	2.1
148	Major small & large bowel procedures w cc	4.5	2.9
107	Coronary bypass w/o cardiac cath	6.6	3.4
14	Specific cerebrovascular disorders except tia	8.4	4.6
106	Coronary bypass w cardiac cath	10.0	4.9
89	Simple pneumonia & pleurisy age > 17 w cc	11.6	6.1
121	Circulatory disorders w ami & c.v. comp disch alive	13.0	6.8
110	Major reconstructive vascular proc w/o pump w cc	14.1	7.1
112	Vascular procedures except major reconstruction w/o pump	15.2	7.7
154	Stomach, esophageal & duodenal procedures age > 17 w cc	16.2	8.2
105	Cardiac valve procedure w pump & w/o cardiac cath	17.1	8.3
108	Other cardiothoracic or vascular procedures, w pump	18.0	8.5
210	Hip & femur procedures except major joint age >17 w cc	18.9	8.8
182	Esophagitis, gastroent & misc digest disorders age >17 w cc	19.9	10.2
79	Respiratory infections & inflammations age >17 w cc	20.7	10.5
96	Bronchitis & asthma age >17 w cc	21.4	11.3
138	Cardiac arrhythmia & conduction disorders w cc	22.2	12.1
416	Septicemia age >17	22.9	12.5
1	Craniotomy age >17 except for trauma	23.6	12.7
296	Nutritional & misc metabolic disorders age > 17 w cc	24.2	13.3
75	Major chest procedures	24.9	13.6
88	Chronic obstructive pulmonary disease	25.4	14.1
197	Total cholecystectomy w/o c.d.e. w cc	26.0	14.6
174	G.I. hemorrhage w cc	26.6	15.1
320	Kidney & urinary tract infections age > 17 w cc	27.1	15.7

NOTE: For complete diagnosis-related group descriptions, see Health Systems International (1988).

SOURCE: Calculated from total 1986 New Jersey Department of Health hospital claims file.

currently allows a 2-percent operating margin in its cost-based payments. It is interesting to note the considerable amount of disagreement between refined payment policies.

A careful evaluation of payment changes and hospital characteristics was performed to correlate different types of hospitals with beneficial or detrimental payment shifts. No simple explanation of the relationship between study hospital characteristics and the payment impact resulting from severity adjustment was found.[11]

Adjacent diagnosis-related group class structures

An indepth analysis of the Yale and CSI class models is now used to compare those models with selected other class models using a common subset of the identified DRGs. The most significant 25 DRGs of the DRG subset, ordered by their percent of total cost, are illustrated in Table 5. As is evident from Table 5, the top 25 DRGs represent 27.1 percent of cost and 15.7 percent of patients.

Six models are summarized in Table 6 for the 25 ADRGs that incorporate the 25 DRGs shown in Table 5. Data for the table are calculated using trimmed study records that exclude records with a discharge status of death or transfer to another acute care facility. The model numbers are as previously given in Table 4, and the number of groups (#Group) indicates the total

number of distinct patient classes formed. For example, the 87 groups formed using Yale Model 8 indicate that the three-class medical and four-class surgical overlays on 25 ADRGs created a total of 87 patient groups. Cost reduction in variance (RIV) for the 87 group Yale Model 8 refinement is 54.4 percent as compared with a cost RIV of 45.7 percent for the unrefined 25 ADRGs. The Yale two-class refinement using Model 10 on medical ADRGs and Model 12 on surgical ADRGs yielded a total of 49 groups; one potential group had no cases. The reduction in groups from 87 to 49 resulted in a cost RIV decrease from 54.4 percent to 52.3 percent. Results for CC Model 15 show the status quo policy, splitting an ADRG on the existence of a CC. It is clear that 72 percent of patients in these ADRGs have CCs and that the payment differential is considerably lower than the two-class CSI and Yale models, where membership in Class 2 is based on the highest CSI or Yale scores only.

Overall, severity refinement Model 1 achieved the greatest improvement in within-class homogeneity, where nearly 60 percent of records were rated severity level 1, with an average cost of $1,925. Complexity refinement Model 8 rated 47 percent of records as level 2 with an average cost of $2,154. The two-class model for CSI and Yale, with 94 percent and 85 percent, respectively, of cases in the lower cost group, achieved comparable results to those achieved by Yale Model 8. The two-class status quo policy, Model 15, is clearly an inferior policy. Optimal Model 21 indicates that a cost RIV of 81.2 percent is an upper bound for any two-class overlay structure.

Data for the table of comparisons are calculated from 20,597 study records. There were 8 percent deaths and

---

[11] A detailed discussion of the cost simulation is included in the project final report. Payment changes as a function of size, Medicaid share, and teaching status were studied and are commented on in the final report.

Table 6
Distribution of severity levels for the top 25 adjacent diagnosis-related groups (ADRGs)

Model	#Group	Cost RIV		Class 1	Class 2	Class 3	Class 4
CSI Model 1	100	58.6	% Count	59	25	10	6
			avg(cost)	$1,925	$3,088	$4,394	$7,631
			% Cost	41	28	16	15
Yale Model 8	87	54.4	% Count	28	47	20	5
			avg(cost)	$1,741	$2,154	$4,450	$8,070
			% Cost	17	36	32	14
CSI Model 5	50	53.3	% Count	94	6	—	—
			avg(cost)	$2,506	$7,631	—	—
			% Cost	85	15	—	—
Yale Models 10, 12	49	52.3	% Count	85	15	—	—
			avg(cost)	$2,358	$5,190	—	—
			% Cost	72	28	—	—
CC Model 15	50	49.1	% Count	28	72	—	—
			avg(cost)	$1,772	$3,184	—	—
			% Cost	18	82	—	—
Optimal Model 21	50	81.2	% Count	81	19	—	—
			avg(cost)	$1,954	$5,466	—	—
			% Cost	60	40	—	—

NOTES: Cost variance reduction using 25 ADRGs for the 20,597 records is 45.7 percent. Group is total number of classes in the 25ADRGs. Cost RIV is the percent reduction in variance of cost. % Count is percent of patients in class. Avg (cost) is average cost of patients in class. % Cost is percent total cost of patients in class.

SOURCE: Calculated from 1988 New Jersey Severity of Illness Evaluation Project data.

Table 7
Cost summary and trim statistics for the top 25 adjacent diagnosis-related groups (ADRGs)

ADRG	N	Mean	Med	CV	Min	Max	Q1	Q3	Low	High	Trim	tpct
1	195	6,251	4,545	0.86	416	35,218	2,988	7,241	792	27,311	6	3.1
14	1,152	3,116	2,183	1.12	184	39,045	1,406	3,462	364	13,385	33	2.9
75	271	5,005	4,056	0.82	234	37,708	2,645	5,918	790	19,808	11	4.1
79	378	3,967	2,692	1.15	278	49,381	1,574	4,499	325	21,754	7	1.9
88	565	2,250	1,701	0.96	235	22,850	1,034	2,700	245	11,385	5	0.9
89	1,521	2,238	1,636	0.96	189	24,029	1,067	2,546	289	9,388	37	2.4
96	1,614	1,600	1,206	0.91	150	31,352	801	1,968	208	7,584	12	0.7
105	116	10,923	9,551	0.51	2,687	37,561	7,260	12,899	3,066	30,545	4	3.4
106	287	9,748	8,711	0.43	3,672	41,837	7,116	11,238	3,586	22,299	6	2.1
107	426	7,039	6,416	0.36	2,237	21,991	5,579	7,929	3,293	13,435	16	3.8
108	159	9,232	8,354	0.52	315	26,382	5,999	11,448	2,276	30,176	3	1.9
110	392	5,707	4,548	0.73	475	39,371	3,343	6,595	1,207	18,270	8	2.0
112	582	3,335	2,912	0.69	272	24,942	1,875	4,072	586	13,034	27	4.6
121	1,595	3,086	2,708	0.75	119	28,176	1,832	3,720	633	10,763	88	5.5
127	1,994	2,155	1,677	0.80	170	18,831	1,116	2,600	314	9,241	29	1.5
138	1,299	1,495	1,196	0.93	102	21,683	724	1,831	180	7,363	17	1.3
148	852	4,814	3,468	0.91	352	46,693	2,575	5,561	812	17,648	24	2.8
154	560	3,616	2,220	1.18	137	46,630	1,285	4,480	197	29,161	4	0.7
174	690	1,673	1,303	0.78	152	13,532	907	2,006	276	6,599	14	2.0
182	2,302	1,190	908	1.03	118	31,265	577	1,432	148	5,597	30	1.3
197	1,131	2,010	1,643	0.69	116	19,070	1,262	2,237	535	5,276	30	2.7
210	531	4,304	3,584	0.64	649	29,814	2,715	4,911	1,116	11,948	19	3.6
296	769	1,989	1,323	1.24	137	38,120	831	2,322	178	10,854	13	1.7
320	872	1,685	1,260	1.00	118	26,498	812	2,039	204	8,111	13	1.5
416	344	3,203	2,541	0.83	159	24,699	1,658	3,831	472	13,450	10	2.9

NOTE: N is the count of records used for each ADRG. Mean is the average cost. Med is the median cost. CV is the coefficient of variation. Min and Max are the extreme cost values. Q1 and Q3 are the first and third cost quartiles. Low and high are cost-trim values using the logarithmically transformed formulas stated earlier. Trim and tpct are the number and percent of records that are excluded using the low- and high-trim values.

SOURCE: Calculated from 1988 New Jersey Severity of Illness Evaluation Project data.

2 percent acute transfer records excluded. There were also 466 records excluded based on the statistical trimming algorithm. Statistics for the untrimmed records for each ADRG are included in Table 7. The last two columns, labeled "trim" and "tpct," yield the number of records that were statistically trimmed and the percent of records trimmed. Relative standard error of cost is the coefficient of variation (CV) divided by the square root of the count (N); the relative standard error for each ADRG is no greater than 6.1 percent.

Conclusion

The evaluation project was undertaken to determine if a state-of-the-art clinical severity system could be implemented in New Jersey hospitals and to yield data for cost-benefit impact analyses. This discussion has focused on the organization of the study, models for class-structure analysis, and likely payment shifts resulting from DRG class refinement. A more detailed comparative

evaluation concerning DRG class structure overlays is available (McGuire, 1989a). The use of severity information in the New Jersey payment system is still being evaluated, and benefit issues include an increased focus on quality of care considerations.

Beginning in January 1990, New Jersey adopted the NY version 7.0 DRG revision that included 55 new major CC DRGs. Major CC DRGs distinguish high-cost surgical and medical patients in each MDC by the existence of a particularly hazardous CC. The addition of major CC DRGs and other New York DRG changes that identify high-cost patients—tracheostomy, multiple trauma, problems related to the human immunodeficiency virus, bone marrow transplants, and other transplants—represent a considerable improvement to Medicare DRGs and have complicated the consideration of the merits of severity class refinement (McGuire, 1989b). Major CC DRGs were formed as a result of this research, where it was shown that identification of high complexity and severity patients had the greatest impact for refining DRGs with the fewest additional groups. Still, the question remains of whether the existence of coded CCs is sufficient to indicate higher payments.

Current payment policy reflects the cost of a patient complication or procedure without attempting to determine if the complication was avoidable or the procedure was necessary. Refinement of DRGs into complexity classes is an extension of this strategy but carries with it increased concern about information accuracy. The impact of incorrect payment resulting from faulty coding is greater when a class structure is overlayed on DRGs. The following example indicates that the additional cost of verifying CCs, perhaps using severity information, may be warranted in certain DRGs.

The ratio of cost weights between the high- and low-cost classes is typically between three and four and indicates that the average patient cost for the highest class is several times the average patient cost for the lowest class. The product of the high-class cost weight and the HCFA DRG 148 cost weight produces an effective cost weight over 6 (1.9 * 3.27 = 6.21); the effective cost weight for the lowest class would be about 2 (0.6 * 3.27 = 1.96). Using the data from Table 3, the low- to high-cost range for patients in ADRG 148 was $3,500 to $10,600; the average cost for all ADRG 148 patients was $5,600.

A single mistaken ICD-9-CM diagnosis code can induce a significant payment shift through a DRG class change, whereas no single error in a physiological finding can induce a severity class change.

Final evaluation of the impact of patient severity on payment is not possible until exogenous issues affecting payment equity, such as physician and nursing practice variability, teaching program costs, and hospital peer group homogeneity, are resolved. However, it is our opinion that a discriminating use of severity measures will improve the equity of the prospective payment system and also provide information for insuring the fairness and appropriateness of payments.

References

American College of Surgeons: *Socio-Economic Factbook for Surgery, 1988*. Surgical Practice Department, American College of Surgeons, Chicago, 1985.

American Hospital Association: *Hospital Statistics*. Chicago, 1987.

Andrews, D. F., Bickel, Hampel, et al.: *Robust Estimates of Location*. Princeton, N.J. Princeton University Press, 1972.

Averill, R. B., McGuire, T. E., Manning, B. E., et al.: *A Study of the Relationship Between Severity of Illness and Hospital Cost in New Jersey Hospitals*. New Haven, Conn. Health Systems International, 1989.

Fetter, R. B., Freeman, J., Park, H., et al.: *DRG Refinement with Diagnosis Specific Comorbidity: A Synthesis of Current Approaches to Patient Classification:* Final Report. Contract Nos. 15-C-98930/1-01 and 17-C-98930/1-0251. Prepared for Health Care Financing Administration Health Systems Management Group. Yale University, New Haven. 1989.

Health Systems International: *DRGs, Diagnosis Related Groups, Fifth Revision, Definitions Manual*. New Haven, Conn. Health Systems International, Appendix G., 1988.

Horn, S. D.: Measuring severity: How sick is sick? How well is well?, *Healthcare Financial Management* Oct. 1986.

Iezonni, L., Moskowitz, M. A., and Daley, J.: *A Description and Clinical Assessment of the Computerized Severity Index*. Boston University Medical Center, Boston, Mass. 1989.

McGuire, T. E.: Evaluation Criteria for Case-Mix Analysis Systems. Unpublished doctoral dissertation. New Haven, Conn. Yale University, 1989a.

McGuire, T. E.: Recent DRG refinements. *Proceedings of the 5th Patient Classification Systems/Europe Conference*. Sabaudia, Italy, Sept. 25-27, 1989b.

McGuire, T. E., and Bender, J. A.: Optimal Variance Reduction. Working paper, New Haven, Conn. Yale University, 1990.

Network Inc.: *Graduate Medical Education Related Direct Patient Care Costs in the New Jersey Hospital Reimbursement Context*. New Jersey, New Jersey Department of Health, 1989.